問7.1

$D = G_1 \cup G_2 \cdots \cup G_r$ かつ $G_i \cap G_j \neq \phi \ (i \neq j)$ より

$1 = P(G_1) + P(G_2) + \cdots + P(G_r)$ をみたし、$P(G_i)P(D|G_i) = P(D \cap G_i)$

より

$$P(G_1)P(D|G_1) + P(G_2)P(D|G_2) + \cdots + P(G_r)P(D|G_r)$$

$$= P(D \cap G_1) + P(D \cap G_2) + \cdots + P(D \cap G_r)$$

$$= P(G_1 + G_2 + \cdots + G_r)$$

$$= P(D) \blacksquare$$

問7.2

1.1の結果を使って

$$P(G_i | D) = \frac{P(G_i \cap D)}{P(D)} = \frac{P(G_i)P(D|G_i)}{\sum_{j=1}^{r} P(G_j)P(D|G_j)} \quad i = 1, 2, \cdots r$$

問7.3

$P(D|G_i)$は観測データが群 i に属すると仮定したときの相対的な出現確率

だから $P(D|G_1) = \dfrac{f(x|\hat{\theta}_1)}{f(x|\hat{\theta}_1) + f(x|\hat{\theta}_2)}$ と表せる。したがって

$$P(G_1|D) = \frac{P(G_1)P(D|G_1)}{P(G_1)P(D|G_1) + P(G_2)P(D|G_2)}$$

$$= \frac{P(G_1) \times \frac{f(x|\hat{\theta}_1)}{f(x|\hat{\theta}_1) + f(x|\hat{\theta}_2)}}{P(G_1) \times \frac{f(x|\hat{\theta}_1)}{f(x|\hat{\theta}_1) + f(x|\hat{\theta}_2)} + P(G_2) \times \frac{f(x|\hat{\theta}_2)}{f(x|\hat{\theta}_1) + f(x|\hat{\theta}_2)}}$$

$$= \frac{P(G_1) f(x|\hat{\theta}_1)}{P(G_1)f(x|\hat{\theta}_1) + P(G_2)f(x|\hat{\theta}_2)} \quad //$$

問7.4

(1) 問7.3の結果を用いると、

$$\frac{p(G_1|x_0)}{p(G_2|x_0)} = \frac{p(G_1)\, f(x_0|\bar{x}_1, S)}{p(G_2)\, f(x_0|\bar{x}_2, S)}$$

(テキストでは事前確率が等しいとしていたので以下 $p(G_1) = p(G_2)$ とする)

また、

$$\frac{p(G_1|x_0)}{p(G_2|x_0)} \begin{cases} \geq 1 \Rightarrow x_0 \in G_1 \\ < 1 \Rightarrow x_0 \in G_2 \end{cases}$$ の判別方式と

$$h(x_0) = \log \frac{f(x_0|\bar{x}_1, S)}{f(x_0|\bar{x}_2, S)} \begin{cases} \geq 0 \\ < 0 \end{cases}$$ の判別方式は同値

したがって $h(x_0)$ の符号を群に対応させる

(2)

$$h(x) = \log \frac{f(x|\bar{x}_1, S)}{f(x|\bar{x}_2, S)}$$

$$= (x-\bar{x}_2)^T S^{-1} (x-\bar{x}_2) - (x-\bar{x}_1)^T S^{-1} (x-\bar{x}_1) \quad \text{—— ①}$$

$$= (\bar{x}_1-\bar{x}_2)^T S^{-1} x - \frac{1}{2}(\bar{x}_1^T S^{-1} \bar{x}_1 - \bar{x}_2 S^{-1} \bar{x}_2)$$

これが $h(x) = w^T x + c$ に等しい。よって、

$$\begin{cases} w = S^{-1}(\bar{x}_1 - \bar{x}_2) \quad (\text{or } w^T = (\bar{x}_1-\bar{x}_2)^T S^{-1}) \\ c = -\frac{1}{2}(\bar{x}_1^T S^{-1} \bar{x}_1 - \bar{x}_2 S^{-1} \bar{x}_2) \quad // \end{cases}$$

問7.5

問7.4の結果を用いると、群 i, j との間の事後確率の比 (問7.4 ①式より $(x-\bar{x}_j) S^{-1} (x-\bar{x}_j) - (x-\bar{x}_i) S^{-1} (x-\bar{x}_i)$) の大小を比べることでその符号から判別する方法は、マハラノビス距離の大小を比べることと等しい。

問7.6

(1) $P(G_1|x) + P(G_2|x) = 1$ を用いて、$\log \dfrac{P(G_1|x)}{P(G_2|x)} = \log \dfrac{P(G_1|x)}{1-P(G_1|x)} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

を $P(G_1|x)$ について解けばよい。$P(G_2|x)$ についても同様。

(2) $P(G_1|x) \longrightarrow P(Y=1|x) = \pi(x)$　と対応付けて、Yはベルヌーイ分布に
$P(G_2|x) \longrightarrow P(Y=0|x) = 1-\pi(x)$　従うので

$$f(y|\pi(x)) = \pi(x)^y \{1-\pi(x)\}^{1-y}, \quad y=0,1$$

ただし

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2)}{1+\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2)}$$

尤度関数は

$$L(\beta_0, \beta_1, \beta_2) = \prod_{i=1}^{n} \pi_i^{y_i} (1-\pi_i)^{1-y_i} \qquad y_i = 0,1$$

対数尤度関数は

$$\ell(\ \ 〃\ \ ) = \sum_{i=1}^{n} y_i(\beta_0 + \beta_1 x_1 + \beta_2 x_2) - \sum_{i=1}^{n} \log\{1+\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2)\} \quad //$$

問7.4

事後確率の比の対数を非線形関数の線形結合で表す

$$\log \frac{P(G_1|x)}{P(G_2|x)} = w^T b(x)$$

$P(G_1|x) + P(G_2|x) = 1$ より $P(G_1|x) = \exp\{w^T b(x)\} / 1+\exp\{w^T b(x)\}$ と表され
問7.3と同様ラベル変数を導入すればYはベルヌーイ分布に従って、対数尤度関数が得られ、数理最適化法で$\hat{w}$が得られる。

さいごに $\log \dfrac{P(G_1|x)}{P(G_2|x)} = \hat{w}\, b(x) \begin{cases} \geq 0 \Rightarrow G_1 \\ < 0 \Rightarrow G_2 \end{cases}$ と判別する。